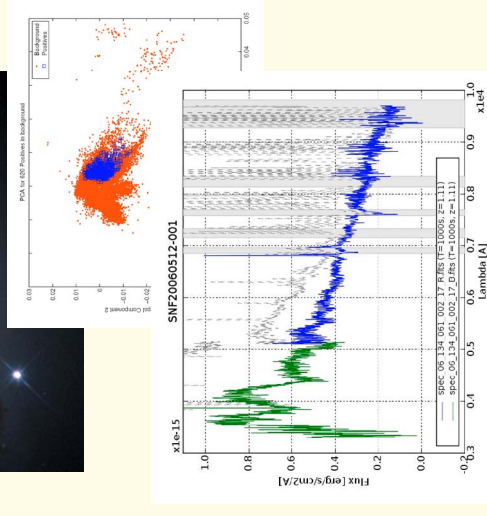
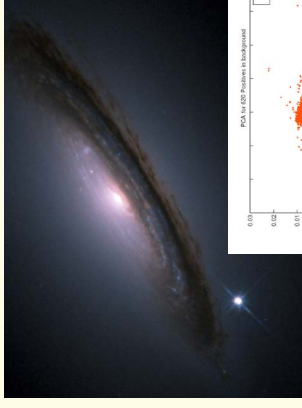




# SpectraVis: Visualization and Analytics for Supernova Spectra

Cecilia R. Aragon, Ali Pinar, Daniela Ushizima  
Lawrence Berkeley National Laboratory  
VACET Meeting, IEEE Vis 2007  
30 October 2007





## Data Deluge in Observational Astronomy

- The Nearby Supernova Factory (SNfactory) is the largest data volume supernova search currently in operation, processing 50-80 GB of image data per night
- Supernova spectral dataset of past 2-3 years is as large as worldwide supernova dataset from past 20 years
- At the same time, computational astrophysicists are developing much more detailed models of supernova progenitors and the process of stellar explosion to attempt to explain the observations

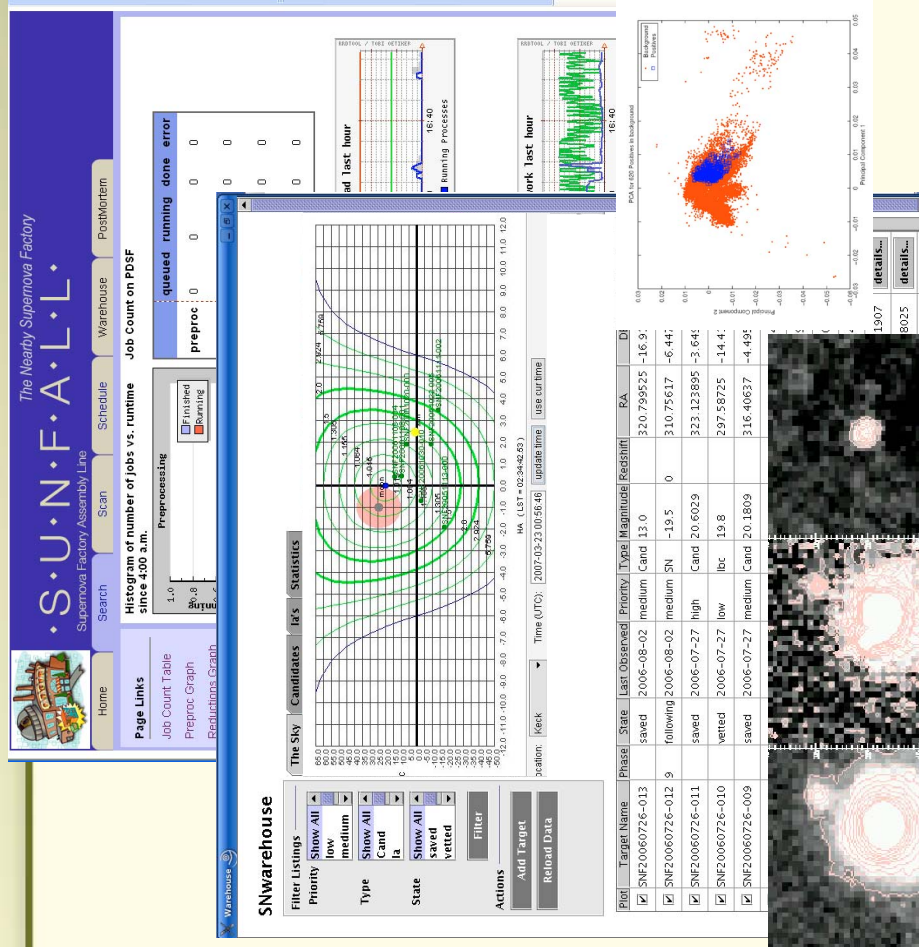


## Three open questions:

1. How can supernova scientists effectively study this large observational data set?
2. Are there spectral features which can encode key supernova attributes?
3. How can we compare model to experimental data?

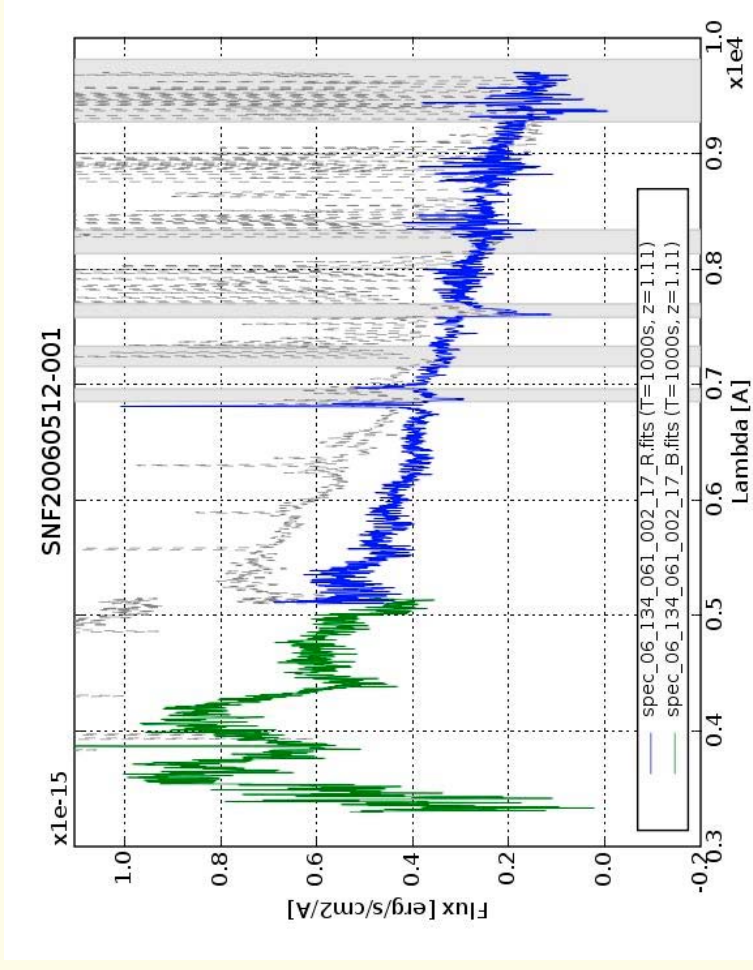
# Leverage Existing Tool: Sunfall

- Sunfall is a collaborative visual analytics system for observational astronomy
- Supernova Warehouse is a graphical interface that provides easy access to all SNfactory spectral and photometric data

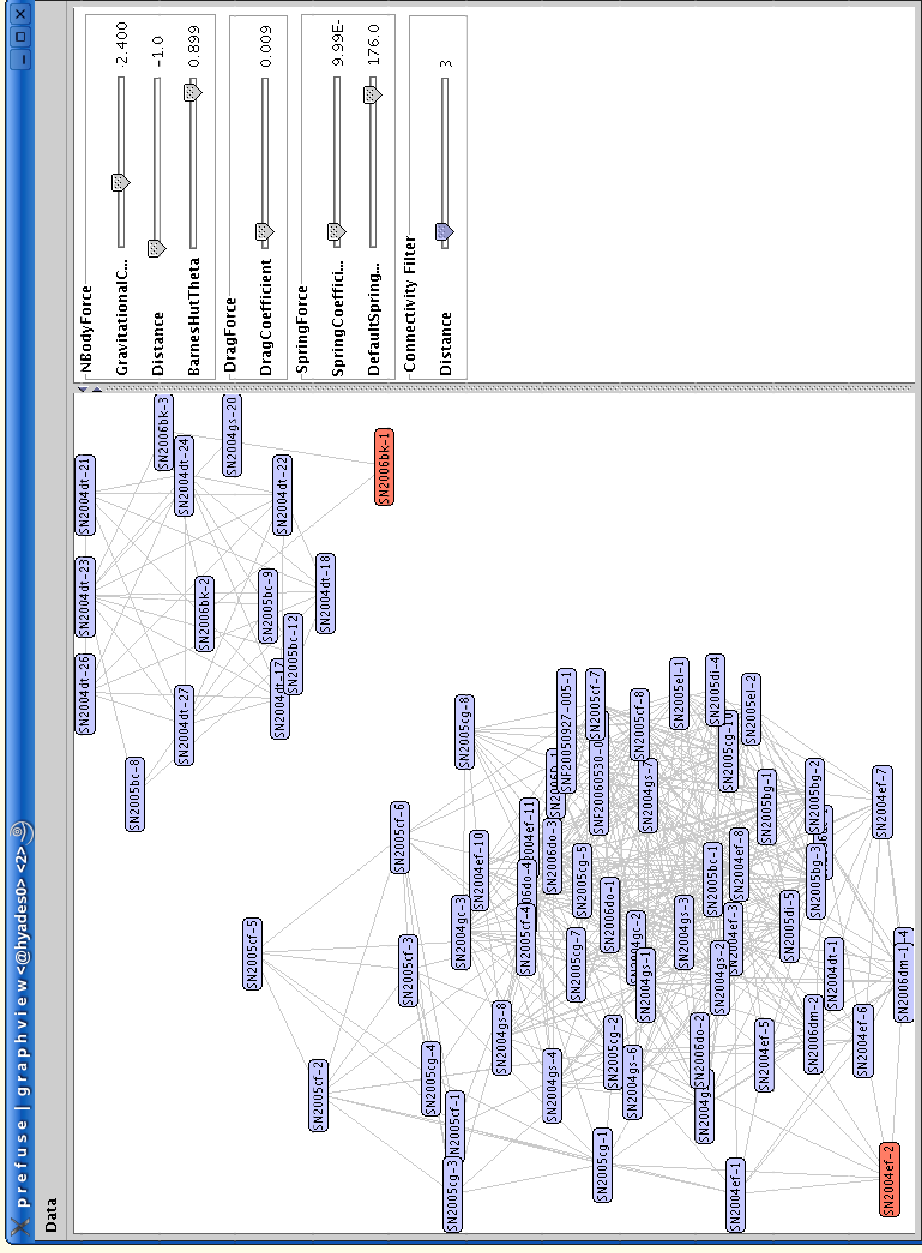




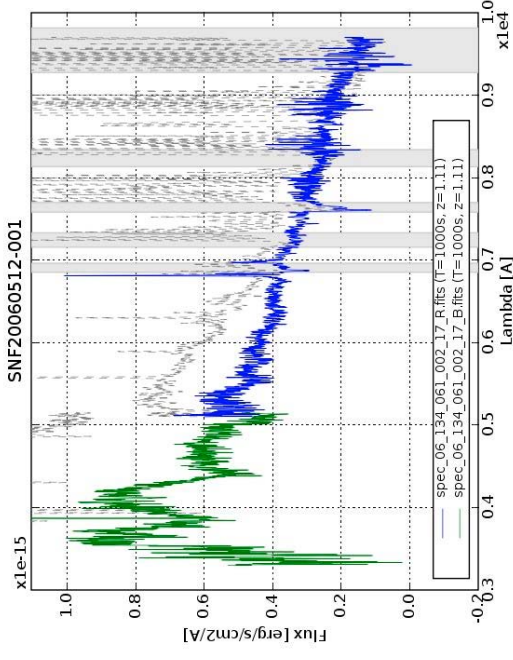
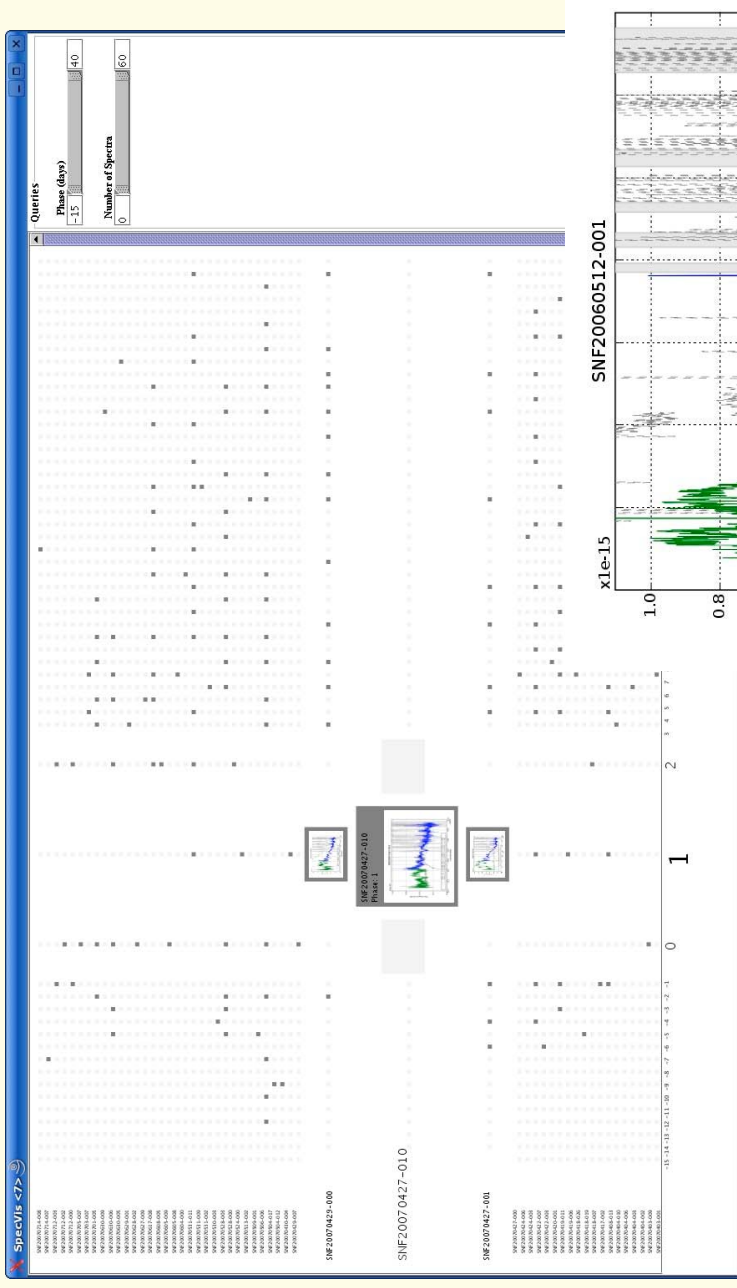
- Need to classify supernova spectra
- Scientists are interested in similarities and outliers
- Need to organize data
  - by date
  - by age
  - by luminosity
  - by type
  - other?



- Clustering using *prefuse*, a Java toolkit
- Dimensionality reduction of supernova spectra
- Similarity measure based on PCA and NMF
- Work by R. Romano, S. Poon, C. Aragon



- 2nd prototype based on scientists' feedback
- Piccolo toolkit for zoomable user interfaces
- Overview grid of entire dataset
- Scientists can zoom and filter
- Work by E. Caraba, S. Poon,



- Calibrated dataset expected November 15
- Previous spectral datasets had systematic errors that interfered with the effectiveness of machine learning techniques
- Three thrusts:
  - Dimensionality reduction for spectral data clustering to detect similarities and outliers - PCA, ICA, NMF, etc.
  - Machine learning techniques to detect set of features which could be used to create filter set for future telescopes
  - Comparisons between model and experiment





# Questions?

Cecilia R. Aragon  
LBNL  
[aragon@hpcrd.lbl.gov](mailto:aragon@hpcrd.lbl.gov)

